

## PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2002-344805

**(43)Date of publication of application : 29.11.2002**

H04N 5/278

H04N 7/025

**H04N 7/03**

H04N 7/035

**(71)Applicant : TELECOMMUNICATION**

**ADVANCEMENT ORGANIZATION  
OF JAPAN**

**NHK ENGINEERING SERVICES INC**

NIPPON HOSO KYOKAI <NHK>

(72)Inventor : SAWAMURA EIJI

**MONMA TAKAO**

**EBARA TERUMASA**

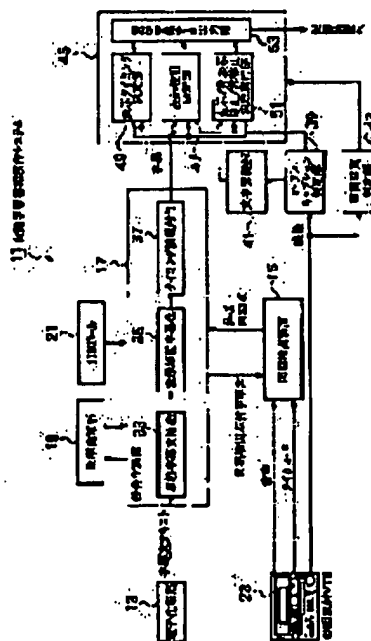
SHIRAI KATSUHIKO

(54) METHOD FOR CONTROLLING SUBTTLES DISPLAY FOR OPEN CAPTION

**(57)Abstract:**

**PROBLEM TO BE SOLVED:** To automatically control the display position and the display contents of display schedule subtitles so as not to cause superimposition when an open caption exists at the display schedule position of the display schedule subtitles which are produced in an automatic subtitles program production system.

**SOLUTION:** An open caption judging part 39 detects the open caption comprised in a video signal to be inputted from a D-VTR 23 through the use of the feature of the open caption which is displayed in a video by superimposition. A character recognizing part 41 converts the open caption detected by the open caption judging part 39 into a text. A screen change judging part 43 detects the changeover point of a screen concerning the video signal to be inputted from the D-VTR 23. A subtitles display position control part 45 performs control to permit the display position of the display schedule subtitles from an integrating device 17 and the display contents not to be superimposed on the open caption.



## LEGAL STATUS

**[Date of request for examination]**

**[Date of sending the examiner's decision of rejection]**

**[Kind of final disposal of application other than the examiner's decision of rejection or**

application converted registration]  
[Date of final disposal for application]  
[Patent number]  
[Date of registration]  
[Number of appeal against examiner's decision  
of rejection]  
[Date of requesting appeal against examiner's  
decision of rejection]  
[Date of extinction of right]

Copyright (C); 1998,2003 Japan Patent Office

(12) 公開特許公報 (A)

(11)特許出願公開番号  
特開2002-344805  
(P2002-344805A)

(43)公開日 平成14年11月29日(2002.11.29)

(51)Int.Cl.	識別記号	F I	テ-マ-ト*(参考)
H 0 4 N	5/278	H 0 4 N	5 C 0 2 3
	7/025		A 5 C 0 6 3
	7/03		
	7/035		

審査請求 未請求 請求項の数 3 OL (全 14 頁)

(21)出願番号	特願2001-148426(P2001-148426)	(71)出願人	592256623 通信・放送機構 東京都港区芝2-31-19
(22)出願日	平成13年5月17日(2001.5.17)	(71)出願人	591053926 財団法人エヌエイチケイエンジニアリング サービス 東京都渋谷区宇田川町37番18号
		(71)出願人	000004352 日本放送協会 東京都渋谷区神南2丁目2番1号
		(74)代理人	100083806 弁理士 三好 秀和 (外3名)

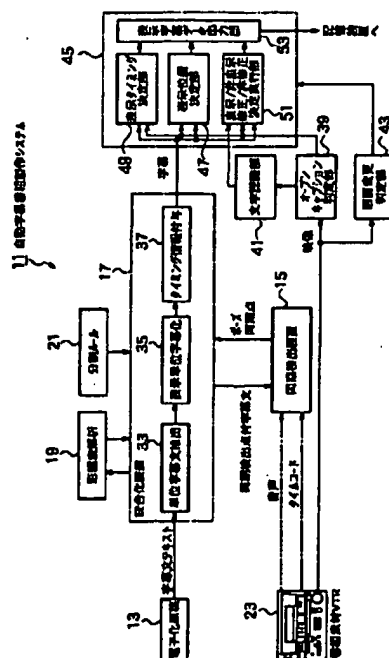
最終頁に続<

(54) 【発明の名称】 オープンキャプションに対する字幕表示制御方法

(57) 【要約】

【課題】 自動字幕番組制作システムにおいて制作した表示予定字幕の表示予定位置に、オープンキャプションが存在するときに、重複しないように表示予定字幕の表示位置や表示内容を自動的に制御できるようにする。

【解決手段】 オープンキャプション判定部３９は、映像にスーパー表示されるオープンキャプションの特徴を利用して、Ｄ－ＶＴＲ２３から入力する映像信号に含まれるオープンキャプションを検出する。文字認識部４１は、オープンキャプション判定部３９で検出できたオープンキャプションをテキスト化する。画面変更判定部４３は、Ｄ－ＶＴＲ２３から入力する映像信号における画面の切り替わり点を検出する。字幕表示位置制御部４５では、統合化装置１７からの表示予定字幕の表示位置が、オープンキャプションと重ならない位置や内容にする制御を行う。



## 【特許請求の範囲】

【請求項1】 テレビ映像にスーパー表示されるオープンキャプションの有無を検知し、自動字幕番組制作システムにおいて制作された表示予定字幕の表示予定位置に前記オープンキャプションが存在するとき、前記表示予定字幕の表示位置を同じフレームにおいて前記オープンキャプションと重ならない所定位置に決定する、または、前記表示予定字幕の表示位置を前記オープンキャプションが存在しないフレームにおける所定位置に決定することを特徴とするオープンキャプションに対する字幕表示制御方法。

【請求項2】 テレビ映像にスーパー表示されるオープンキャプションの有無及びそのオープンキャプションの内容を検知し、自動字幕番組制作システムにおいて制作された表示予定字幕の表示予定位置に前記オープンキャプションが存在するとき、前記オープンキャプションの内容と前記表示予定字幕の内容との近似度を調査し、調査した近似度の程度に応じて、前記表示予定字幕の表示／非表示ないしは修正／非修正を決定することを特徴とするオープンキャプションに対する字幕表示制御方法。

【請求項3】 請求項1又は請求項2に記載のオープンキャプションに対する字幕表示制御方法において、前記テレビ映像のシーン変更点を検出し、検出したシーン変更点と字幕終了タイミングとを比較し、比較結果に基づき字幕終了タイミングを変更することを特徴とするオープンキャプションに対する字幕表示制御方法。

## 【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、自動字幕番組制作システムにおいて制作された表示予定字幕の表示位置やタイミングなどをオープンキャプションに対して自動的に制御する方法に関するものである。

【0002】

【従来の技術】現代は高度情報化社会と一般に言われているが、聴覚障害者は健康者と比較して情報の入手が困難な状況下におかれている。即ち、例えば、情報メディアとして広く普及しているTV放送番組を例示すると、TV放送番組に対する字幕番組の割合は、欧米では33～70%に達しているのに対し、我が国ではわずか10%程度と比べて低くおかれているのが現状である。

【0003】このように、我が国で全TV放送番組に対する字幕番組の割合が欧米と比較して低くおかれている要因としては、主として字幕番組制作技術の未整備を挙げることができる。具体的には、日本語特有の問題もあり、字幕番組制作工程の殆どが手作業によっており、多大な労力・時間・費用を要するためである。

【0004】そこで、本発明者らは、字幕番組制作技術の整備を妨げている原因究明を企図して、現行の字幕番組制作の実態調査を行った。図10の左側には、現在一般に行われている字幕番組制作フローを示してある。図

10の右側には、改良された現行字幕制作フローを示してある。

【0005】図10の左側において、ステップS101では、字幕番組制作者が、タイムコードを映像にスーパーした番組データと、タイムコードを音声チャンネルに記録した番組テープと、番組台本との3つの字幕原稿作成素材を放送局から受け取る。なお、図中において「タイムコード」を「TC」と略記する場合があることを付言しておく。

【0006】ステップS103では、放送関係経験者等の専門家が、ステップS101で受け取った字幕原稿作成素材を基に、(1)番組アナウンスの要約書き起こし、(2)別途規定された字幕提示の基準となる原稿作成要領に従う字幕提示イメージ化、(3)その開始・終了タイムコード記入、の各作業を順次行い、字幕原稿を作成する。

【0007】ステップS105では、入力オペレータが、ステップS103で作成された字幕原稿をもとに電子化字幕データを作成する。ステップS107では、ステップS105で作成された電子化字幕データを、担当の字幕制作責任者、原稿作成者、及び入力オペレータの三者立ち会いのもとで試写・修正を行い、完成字幕とする。

【0008】ところで、最近では、番組アナウンスの要約書き起こしと字幕の電子化双方に通じたキャプションオペレータと呼ばれる人材を養成することで、図10の右側に示す改良された現行字幕制作フローも一部実施されている。

【0009】即ち、ステップS111では、字幕番組制作者が、タイムコードを音声チャンネルに記録した番組テープと、番組台本との2つの字幕原稿作成素材を放送局から受け取る。

【0010】ステップS113では、キャプションオペレータが、タイムコードを音声チャンネルに記録した番組テープを再生する。このとき、セリフの開始点でマウスのボタンをクリックすることでその点の音声チャンネルから始点タイムコードを取り出して記録する。さらに、セリフを聴取して要約電子データとして入力する。同様に、字幕原稿作成要領に基づく区切り箇所に対応するセリフ点で再びマウスのボタンをクリックすることでその点の音声チャンネルから終点タイムコードを取り出して記録する。これらの操作を番組終了まで繰り返して、番組全体の字幕を電子化する。

【0011】ステップS117では、ステップS105で作成された電子化字幕データを、担当の字幕制作責任者、及びキャプションオペレータの二者立ち会いのもとで試写・修正を行い、完成字幕とする。

【0012】後者の改良された現行字幕制作フローでは、キャプションオペレータが、タイムコードを音声チャンネルに記録した番組テープのみを使用して、セリフ

の要約と電子データ化を行うとともに、提示単位に分割した字幕の始点/終点にそれぞれ対応するセリフのタイミングでマウスボタンをクリックすることにより、音声チャンネルの各タイムコードを取り出して記録するものであり、かなり省力化された効果的な字幕制作フローといえる。

【0013】ここで、上述した現行字幕制作フローにおける一連の処理の流れの中で特に多大な工数を要するのは、ステップS103ないしはS105またはステップS113の、(1)番組アナウンスの要約書き起こし、(2)字幕提示イメージ化、(3)その開始・終了タイムコード記入、の各作業工程である。これらの作業工程は熟練者の知識・経験に負うところが大きい。

【0014】ところが、現在放送中の字幕番組の中で、予めアナウンス原稿が作成され、その原稿が殆ど修正されることなく実際の放送字幕となっていると推測される番組がいくつかある。例えば、「生きもの地球紀行」という字幕付き情報番組を実際に調べて見ると、アナウンス音声と字幕内容は殆ど共通であり、共通の原稿をアナウンス用と字幕用の双方に利用しているものと推測できる。

【0015】そこで、本発明者らは、ほぼ共通の電子化原稿をアナウンス用と字幕用の双方に利用する形態を想定して、字幕番組の制作を自動化できる自動字幕番組制作システムを開発し先に出願した(例えば特開2000-270623号公報)。

【0016】

【発明が解決しようとする課題】しかし、テレビ映像にスーパー表示される文字、いわゆるオープンキャプションが表示されているときに、映像上に表示しようとする字幕放送の字幕がオープンキャプションと重なる場合には、双方とも非常に見難くなる。

【0017】そのため、映像のオープンキャプションを調べて、それと重複しない表示位置やタイミングを字幕用原稿の段階で指定するとか、試写・修正段階で表示位置やタイミングを重複しないよう修正するなどして、オープンキャプションに対し字幕が重ならないようにする作業が必要になる。これらの作業は全て専門知識を有する人が手作業により、多くの時間を使って行われていた。

【0018】本発明は、このような実情に鑑みてなされたものであり、ほぼ共通の電子化原稿をアナウンス用と字幕用の双方に利用する形態を想定して字幕番組を制作する自動字幕番組制作システムにおいて制作した表示予定字幕の表示予定位置に、オープンキャプションが存在するときに、重複しないように表示予定字幕の表示位置やタイミングを自動的に制御できるオープンキャプションに対する字幕表示制御方法を提供することを目的としている。

【0019】

【課題を解決するための手段】上記目的を達成するために、請求項1に記載のオープンキャプションに対する字幕表示制御方法は、テレビ映像にスーパー表示されるオープンキャプションの有無を検知し、自動字幕番組制作システムにおいて制作された表示予定字幕の表示予定位置に前記オープンキャプションが存在するとき、前記表示予定字幕の表示位置を同じフレームにおいて前記オープンキャプションと重ならない所定位置に決定する、または、前記表示予定字幕の表示位置を前記オープンキャプションが存在しないフレームにおける所定位置に決定することを特徴とする。

【0020】この方法によれば、ほぼ共通の電子化原稿をアナウンス用と字幕用の双方に利用する形態を想定して字幕番組を制作する自動字幕番組制作システムにおける字幕制作の一環として、テレビ映像にスーパー表示されるオープンキャプションの有無を検知し、表示予定の字幕が検知したオープンキャプションと重ならないようにその字幕の表示位置や表示タイミングをオープンキャプションに応じて自動的に制御できる。

【0021】請求項2に記載のオープンキャプションに対する字幕表示制御方法は、テレビ映像にスーパー表示されるオープンキャプションの有無及びそのオープンキャプションの内容を検知し、自動字幕番組制作システムにおいて制作された表示予定字幕の表示予定位置に前記オープンキャプションが存在するとき、前記オープンキャプションの内容と前記表示予定字幕の内容との近似度を調査し、調査した近似度の程度に応じて、前記表示予定字幕の表示/非表示ないしは修正/非修正を決定することを特徴とする。

【0022】この方法によれば、オープンキャプションをテキストとしても自動検知し、表示予定の字幕内容と比較して、字幕の表示/非表示、字幕内容の修正/非修正なども自動的に判別し実行することができる。

【0023】請求項3に記載のオープンキャプションに対する字幕表示制御方法は、請求項1又は請求項2に記載のオープンキャプションに対する字幕表示制御方法において、前記テレビ映像のシーン変更点を検出し、検出したシーン変更点と字幕終了タイミングとを比較し、比較結果に基づき字幕終了タイミングを変更することを特徴とする。

【0024】この方法によれば、例えば、カットのタイミングを遅けるように字幕の開始終了のタイミングを制御することができ、映像・音声と字幕が一致しないような不適切な字幕表示を無くすることができる。

【0025】

【発明の実施の形態】以下、本発明の実施の形態について、添付図面を参照して詳細に説明する。

【0026】図1は、本発明に係るオープンキャプションに対する字幕表示制御方法を実施する自動字幕番組制作システムの機能ブロック構成図である。図2は、自動

字幕番組制作システムにおける字幕制作フローを、改良された現行字幕制作フローと対比して示した説明図である。図3は、単位字幕文を表示単位字幕文毎に分割する際に適用される分割ルールの説明に供する図である。図4ないしは図5は、アナウンス音声に対する字幕送出タイミングの同期検出技術に係る説明に供する図である。図6は、本発明の実施の形態1に係るオープンキャプションに対する字幕表示制御方法を説明するフローチャートである。図7は、実施の形態1による字幕の表示位置及び開始・終了のタイミングの変更例を示す図である。図8は、本発明の実施の形態2に係るオープンキャプションに対する字幕表示制御方法を説明するフローチャートである。図9は、実施の形態2による字幕表示の適性化処理例を示す図である。

【0027】図1に示すように、自動字幕番組制作システム11は、電子化原稿記録媒体13と、同期検出装置15と、統合化装置17と、形態素解析部19と、分割ルール記憶部21と、番組素材VTR例えばデジタル・ビデオ・テープ・レコーダ（以下、「D-VTR」と言う）23と、オープンキャプション判定部39と、文字認識部41と、画面変更判定部43と、字幕表示位置制御部45とを含んで構成されている。このうち、オープンキャプション判定部39と文字認識部41と画面変更判定部43と字幕表示位置制御部45とが本実施の形態に係る字幕表示制御を行う機能部分である。

【0028】ここでは、まず、図1～図5を用いて自動字幕番組制作システムにおける字幕制作の概要を説明する。その後で、本実施の形態に係る字幕表示制御を、実施の形態1及び実施の形態2として説明する。

【0029】既述したように、現在放送中の字幕番組の中で、予めアナウンス原稿が作成され、その原稿が殆ど修正されることなく実際の放送字幕となっていると推測される番組がいくつかある。例えば、「生きもの地球紀行」という字幕付き情報番組を実際に調べて見ると、アナウンス音声と字幕内容はほぼ共通であり、ほぼ共通の原稿をアナウンス用と字幕用の両方に利用していると推測できる。

【0030】本発明者らが提案する自動字幕番組制作システムは、アナウンス用と字幕用の両方に共通の原稿が電子化されている番組におけるその電子化原稿に、本発明者らが提案するアナウンス音声と字幕文テキストの同期検出技術、及び日本語の特徴解析手法を用いたテキスト分割技術等を適用することにより、D-VTR23から再生されたアナウンスの進行と同期して、表示単位字幕文の作成、及びその始点/終点の各々に対応するタイミング情報の付与を自動化し、これをもって、字幕番組の制作を人手を介することなく自動化できるようにしたものである。

【0031】ここで、具体的な字幕制作の説明に先立って、その説明で使用する用語の定義付けを行う。即ち、

表示対象となる字幕の全体集合を「字幕文テキスト」と言う。字幕文テキストのうち、句読点で区切られた文章単位の部分集合を「単位字幕文」と言う。ディスプレイの表示画面上における表示単位字幕の全体集合を「表示単位字幕群」と言う。表示単位字幕群のうち、任意の一行の字幕を「表示単位字幕文」と言う。表示単位字幕文のうちの任意の文字を表現するとき、それを「字幕文字」と言うことにする。

【0032】図1において、電子化原稿記録媒体13は、例えばハードディスク記憶装置やフロッピー（登録商標）ディスク装置等より構成され、表示対象となる字幕の全体集合を表す字幕文テキストを記憶している。なお、ここでは、ほぼ共通の電子化原稿をアナウンス用と字幕用の双方に利用する形態を想定しているので、電子化原稿記録媒体13に記憶される字幕文テキストの内容は、表示対象字幕とするばかりでなく、D-VTR23のアナウンス音声とも一致しているものとする。

【0033】同期検出装置15は、表示単位字幕文と、それを読み上げたアナウンス音声との間における時間同期を補助する機能等を有している。具体的には、妥当性検証機能と検証結果返答機能とタイミング情報検出機能とを有している。妥当性検証機能とは、同期検出装置15は、統合化装置17で確定された表示単位字幕配列が送られてくる毎に、その表示単位字幕配列の妥当性を検証する機能である。検証結果返答機能とは、妥当性検証機能を発揮することで得られた検証結果が不当であるとき、その検証結果を統合化装置17宛に返答する機能である。タイミング情報検出機能とは、妥当性検証機能を発揮することで得られた検証結果が妥当であるとき、D-VTR23から取り込んだその表示単位字幕配列に対応するアナウンス音声及びそのタイムコードを参照して、該当する表示単位字幕文毎のタイミング情報、即ち始点/終点タイムコードを検出し、検出した各始点/終点タイムコードを統合化装置17宛に送出する機能である。

【0034】形態素解析部19は、漢字かな交じり文で表記されている単位字幕文を対象として、形態素毎に分割する分割機能と、その分割機能によって分割された各形態素毎に、表現形、品詞、読み、標準表現などの付加情報を付与する付加情報付与機能と、各形態素を文節や節単位にグループ化し、いくつかの情報素列を得る情報素列取得機能とを有している。これにより、単位字幕文は、表面素列、記号素列（品詞列）、標準素列、及び情報素列として表現される。

【0035】分割ルール記憶部21は、図3に示すように、単位字幕文を対象とした改行・改頁箇所の最適化を行う際に参照される分割ルールを記憶している。D-VTR23は、番組素材が収録されている番組素材VTRテープから、映像、音声、及びそれらのタイムコードを再生出力する機能を有している。

【0036】統合化装置17は、単位字幕文抽出部33と、表示単位字幕化部35と、タイミング情報付与部37とを有している。単位字幕文抽出部33は、電子化原稿記録媒体13から読み出した字幕文テキストの中から、例えば40～50字幕文字程度を目安とした単位字幕文を順次抽出する。具体的には、単位字幕文抽出部33は、少なくとも表示単位字幕文よりも多い文字数を呈する表示対象となる単位字幕文を、必要に応じその区切り可能箇所情報等を活用して表示時間順に順次抽出する機能を有している。なお、区切り可能箇所情報として

は、形態素解析部19で得られた文節データ付き形態素解析データ及び分割ルール記憶部21に記憶されている分割ルール（改行・改頁データ）を例示することができる。

【0037】表示単位字幕化部35は、単位字幕文抽出部33で抽出された単位字幕文を、所望の表示形式に従う表示単位字幕文に変換する。具体的には、表示単位字幕化部35は、単位字幕文抽出部33で抽出された単位字幕文、その単位字幕文に付加されている区切り可能箇所情報、及び同期検出装置15からの情報等に基づいて、単位字幕文抽出部33で抽出された単位字幕文を、所望の表示形式に従う少なくとも1以上の表示単位字幕文に変換する機能を有している。

【0038】タイミング情報付与部37は、表示単位字幕化部35で変換された表示単位字幕文に対し、同期検出装置15から送出されてきた表示単位字幕文毎のタイミング情報である始点／終点の各タイムコードを付与することを行う。具体的には、タイミング情報付与部37は、表示単位字幕化部35で変換された表示単位字幕文に対し、同期検出装置15から送出されてきた表示単位字幕文毎のタイミング情報である始点／終点の各タイムコードを付与するタイミング情報付与機能を有している。

【0039】次に、自動字幕番組制作システム11における字幕制作動作について、図2の右側に示す字幕制作フローに従って、図2の左側に示す改良された現行字幕制作フローと対比しつつ説明する。

【0040】まず、図2の左側に示す改良された現行字幕制作フローについて再度説明する。ステップS111において、字幕番組制作者は、音声チャンネルにタイムコードを記録した番組テープと、番組台本との2つの字幕原稿作成素材を放送局から受け取る。なお、図中において「タイムコード」を「TC」と略記する場合があることを付言しておく。

【0041】ステップS113において、キャプションオペレータは、VTRの別の音声チャンネル（セリフをLchとするとRch）にタイムコードを記録した番組テープを再生し、セリフの開始点でマウスのボタンをクリックすることでその点の音声チャンネルから始点タイムコードを取り出して記録する。さらに、セリフを聴取

して要約電子データとして入力するとともに、字幕原稿作成要領に基づいて行う区切り箇所に対応するセリフ点で再びマウスのボタンをクリックすることでその点の音声チャンネルから終点タイムコードを取り出して記録する。これらの操作を番組終了まで繰り返し、番組全体の字幕を電子化する。

【0042】ステップS117において、ステップS105で作成された電子化字幕データを、担当の字幕制作責任者、及びキャプションオペレータの二者立ち会いのもとで試写・修正を行い、完成字幕とする。

【0043】このように、改良された現行字幕制作フローでは、キャプションオペレータは、タイムコードをVTRの別の音声チャンネルに記録した番組テープのみを使用して、セリフの要約と電子データ化を行うとともに、表示単位に分割した字幕の始点／終点にそれぞれ対応するセリフのタイミングでマウスボタンをクリックすることにより、音声チャンネルの各タイムコードを取り出して記録するものであり、かなり省力化された効果的な字幕制作を実現している。

【0044】ところが、本発明らによる字幕制作方法によれば、図2の右側に示す字幕制作フローから理解できるように、さらなる省力化が図られている。ここでの処理は、図2の左側に示すステップS113の要約原稿・電子データ作成処理に相当するものである。

【0045】即ち、ステップS1では、単位字幕文抽出部33は、電子化原稿記録媒体13から読み出した字幕文テキストの中から、40～50文字程度を目安として、少なくとも表示単位字幕文よりも多い文字数を呈する単位字幕文を、その区切り可能箇所情報等を活用して順次抽出する。なお、制作する字幕は、通常一行当たり15文字を限度として、二行の表示単位字幕群を順次入れ換えていく字幕表示形式が採用されるので、文頭から40～50字幕文字程度で、句点や読点を目安にして単位字幕文を抽出する（これは15文字の処理量をも考慮している。）。

【0046】ステップS2ないしはS5では、表示単位字幕化部35は、単位字幕文抽出部33で抽出された単位字幕文、及び単位字幕文に付加された区切り可能箇所情報等に基づいて、単位字幕文抽出部33で抽出された単位字幕文を、所望の表示形式に従う少なくとも1以上の表示単位字幕文に変換する。

【0047】具体的には、単位字幕文抽出部33で抽出された単位字幕文を、上述した字幕表示形式に従い、例えば、一行当たり13字幕文字で、二行の表示単位字幕群となる表示単位字幕配列案を作成する（ステップS2）。他方、単位字幕文抽出部33で抽出された単位字幕文を対象とした形態素解析を行い、形態素解析データを得る（ステップS3）。この形態素解析データには文節を表すデータも付属している。そして、上記のように作成した表示単位字幕配列案に対し、形態素解析データ

10

20

30

40

50

を参照して、表示単位字幕配列案の改行・改頁点を最適化し（ステップS4）、最初の単位字幕文に関する表示単位字幕配列を確定する（ステップS5）。これにより、実情に即して高精度に最適化された表示単位字幕化を実現することができる。

【0048】なお、ステップS4において表示単位字幕配列案を最適化するにあたっては、別途用意した分割ルール（改行・改頁データ）も併せて適用する。具体的には、図3に示すように、分割ルール（改行・改頁データ）で定義される改行・改頁推奨箇所は、第1に句点の後、第2に読点の後、第3に文節と文節の間、第4に形態素品詞の間、を含んでおり、分割ルール（改行・改頁データ）を適用するにあたっては、上述した記述順の先頭から優先的に適用する。このようにすれば、さらに実情に即して高精度に最適化された表示単位字幕化を実現することができる。

【0049】特に、第4の形態素品詞の間を分割ルール（改行・改頁データ）として適用するにあたっては、図3の図表には、自然感のある改行・改頁を行った際における、直前の形態素品詞とその頻度例が示されているが、図3の図表のうち頻度の高い形態素品詞の直後で改行・改頁を行うようにすればよい。このようにすれば、より一層実情に即して高精度に最適化された表示単位字幕化を実現することができる。

【0050】ステップS6ないしはS7では、タイミング情報付与部37は、表示単位字幕化部35で変換された表示単位字幕文に対し、同期検出装置15から送出されてきた表示単位字幕文毎のタイミング情報である始点／終点の各タイムコードを付与する。

【0051】具体的には、統合化装置17は、ステップS5で確定した表示単位字幕文を同期検出装置15に与える。一方、同期検出装置15は、ステップS6でD-VTR23からアナウンス音声及びそのタイムコードを取り込むと、ステップS5で確定した表示単位字幕配列、即ち表示単位字幕文に対応するアナウンス音声中に例えば2秒以上の所定時間を超える非スピーチ区間、即ちポーズの存在有無を調査する（ステップS7）。この調査の結果、アナウンス音声中にポーズ有りを検出したときには、該当する表示単位字幕文は不当であるとみなして、ステップS5の表示単位字幕配列確定処理に戻り、このポーズ以前に対応する単位字幕文の中から、表示単位字幕配列を再変換する。

【0052】他方、同期検出装置15は、上記調査の結果、所定時間を超えるポーズ無しを検出したときには、該当する表示単位字幕文は妥当であるとみなして、その始点／終点タイムコードを検出し（ステップS7）、検出した各始点／終点タイムコードを該当する表示単位字幕文に付与して（ステップS8）、最初の単位字幕文に関する表示単位字幕文の作成処理を終了する。

【0053】ここで、ステップS7において表示単位字

幕文に対応するアナウンス音声中のポーズの有無を調査する趣旨は、次の通りである。即ち、表示単位字幕文中に所定時間を超えるポーズが存在するということは、この表示単位字幕文は、時間的に離れており、また、少なくとも複数の相異なる場面に対応する字幕文を含んで構成されているおそれがあり、これらの字幕文を一つの表示単位字幕文とみなしたのでは好ましくないおそれがあるからである。これにより、ステップS5で一旦確定された表示単位字幕文の妥当性を、対応するアナウンス音声の観点から再検証可能となる結果として、好ましい表示単位字幕文の変換確定に多大な貢献を果たすことができる。

【0054】なお、ステップS7における表示単位字幕文に付与する始点／終点タイムコードの同期検出は、本発明者らが研究開発したアナウンス音声を対象とした音声認識処理を含むアナウンス音声と字幕文テキスト間の同期検出技術を適用することで高精度に実現可能である。以下、その概要を図4、図5を用いて説明する。

【0055】即ち、字幕送出タイミング検出の流れは、図4に示すように、まず、かな漢字交じり文で表記されている字幕文テキストを、音声合成などで用いられている読み付け技術を用いて発音記号列に変換する。この変換には、「日本語読み付けシステム」を用いる。

【0056】次に、予め学習しておいた音響モデル（HMM：隠れマルコフモデル）を参照し、「音声モデル合成システム」によりこれらの発音記号列をワード列ベアモデルと呼ぶ音声モデル（HMM）に変換する。そして、「最尤照合システム」を用いてワード列ベアモデルにアナウンス音声を通して比較照合を行うことにより、字幕送出タイミングの同期検出を行う。字幕送出タイミング検出の用途に用いるアルゴリズム（ワード列ベアモデル）は、キーワードスポッティングの手法を採用している。キーワードスポッティングの手法として、フォワード・バックワードアルゴリズムにより単語の事後確率を求め、その単語尤度のローカルピークを検出する方法が提案されている。

【0057】ワード列ベアモデルは、図5に示すように、これを応用して字幕と音声を同期させたい点、即ち同期点の前後でワード列1（Keywords1）とワード列2（Keywords2）とを連結したモデルになっており、ワード列の midpoint（B）で尤度を観測してそのローカルピークを検出し、ワード列2の発話開始時間を高精度に求めることを目的としている。ワード列は、音素HMMの連結により構成され、ガーベジ（Garbage）部分は全音素HMMの並列な枝として構成されている。

【0058】また、アナウンサーが原稿を読む場合、内容が理解しやすいように息継ぎの位置を任意に定めることから、ワード列1、2間にポーズ（Pause）を挿入している。なお、ポーズ時間の検出に関しては、D-VTR23から音声とそのタイムコードが供給され、例えばそ



の音声レベルが指定レベル以下で連続する開始、終了タイムコードとして、周知の技術で容易に達成できる。そして、第一頁目に関する字幕作成が終了すると、続いて第一頁目の次からの字幕文を抽出して第二頁目の字幕化に進み、同様の処理により当該番組の全字幕化を行う。

【0059】さて、本実施の形態に係るオープンキャプションに対する字幕表示制御方法では、以下の構成により、オープンキャプションの特徴を利用してその表示位置やタイミングを検知し、それと重複しないように字幕の表示位置やタイミングを自動的に制御すること（実施の形態1）と、さらにオープンキャプションをテキストとしても検知し、表示予定の字幕と比較して字幕の表示／非表示、字幕内容の修正／非修正なども自動的に行うこと（実施の形態2）が行えれるようになっている。

【0060】図1において、オープンキャプション判定部39は、映像にスーパー表示されるオープンキャプションの特徴を利用して、D-VTR23から入力する映像信号に含まれるオープンキャプションを検出する。即ち、オープンキャプションは、背景映像と比較すると通常次のような相違がある。（1）スーパーする画面位置が限られている、（2）RGBの振幅レベルが大きい（例えば85%以上である）、（3）通常数秒以上静止している、（4）色は通常明るい白色である（これは表示領域での占有率を定める要素である）、（5）スーパー表示される文字の大きさはほぼ一定である（規定の寸法で一文字はほぼ正方形である）、（6）スーパー表示される文字は通常横または縦方向に複数連なっている（複数行の場合もある）、（7）特定のスーパー文字についてはパターンマッチング法の適用が可能である。

【0061】したがって、この相違をうまく活用することによって、一定の条件下ではオープンキャプションの表示タイミングやその内容を自動検出することが可能である。これらの相違点のうち、（1）～（4）をうまく組み合わせることで、かなりのケースでタイミング検出ができ、場合によっては（5）～（7）の適用でさらに精度が高くできる。

【0062】文字認識部41は、オープンキャプション判定部39で検出できたオープンキャプションをテキスト化する。画面変更判定部43は、D-VTR23から入力する映像信号における画面の切り替わり点を検出する。具体的には、画面変更判定部43では、パン操作、チルト操作、ズームイン操作、ズームアウト操作、ワイプ操作、ディゾルブ操作等がなされた画面の指定操作の開始と終了を検出している。

【0063】字幕表示位置制御部45では、統合化装置17からの字幕（表示予定字幕）が、表示位置決定部47と表示タイミング決定部49と表示／非表示・修正／非修正決定部51とに入力している。また、表示位置決定部47と表示タイミング決定部49と表示／非表示・修正／非修正決定部51とにオープンキャプション判定

部39の判定結果が入力している。一方、表示／非表示・修正／非修正決定部51には、文字認識部41で認識されたオープンキャプションのテキストが入力している。

【0064】そして、表示位置決定部47と表示タイミング決定部49と表示／非表示・修正／非修正決定部51とは、それぞれ、画面変更判定部43から指定された画面期間内において所定の動作を行うようになっている。表示字幕データ出力部53は、表示位置決定部47と表示タイミング決定部49と表示／非表示・修正／非修正決定部51とでそれぞれ決定された表示字幕データをハードディスクやフロッピーディスクなどの記憶装置へ蓄積する。

【0065】表示位置決定部47では、表示予定字幕の表示位置がオープンキャプションが検出された画面の所定領域と重なる場合に、表示予定字幕の表示位置の移動により、表示予定字幕の表示位置を同じフレーム内で重ならない位置、場合によってはオープンキャプションが存在しない他のフレームでの位置を決定し、その決定した位置データを付けた表示予定字幕を表示字幕データ出力部53に渡すことを行う。

【0066】表示タイミング決定部49では、表示予定字幕の表示位置がオープンキャプションが検出された画面の所定領域と重なる場合に、表示予定字幕の表示開始・終了タイミングの変更により同じフレーム内で重ならない位置を決定し、表示予定字幕の表示位置をその決定した位置データを付けた表示予定字幕を表示字幕データ出力部53に渡すことを行う。

【0067】表示／非表示・修正／非修正決定部51では、表示予定字幕の表示位置がオープンキャプションが検出された画面の所定領域と重なる場合に、オープンキャプションの内容と表示予定字幕の内容とを比較し、当該字幕を表示するか否か、当該字幕を修正するか否かを決定し、その決定内容を実行し、実行結果による字幕データを表示字幕データ出力部53に渡すことを行う。

【0068】以下、文字認識部41を用いないで表示位置やタイミング制御を行う場合（実施の形態1）と、文字認識部41を用いて表示制御などを行う場合（実施の形態2）とに分けて説明する。

【0069】（実施の形態1）文字認識部41を用いないで表示位置やタイミング制御を行う場合には、オープンキャプションの検出結果を利用して、これから送出しようとする表示単位の字幕信号を、どの位置に、どのような開始・終了タイミングで送出するかを制御する。制御方法として、（1）字幕の位置のみ、（2）開始・終了タイミングのみ、（3）位置と開始・終了タイミングの双方、の3方法があり、オープンキャプションと表示する字幕との関係に応じて適用する。

【0070】（1）字幕の位置のみの制御では、オープンキャプションと異なる位置に字幕を移動する。これ

は、表示する字幕の開始・終了タイミングを変えたくない場合に行われる。移動する位置は、既に検出済みのオープンキャプションの位置と大きさから決定する。通常、オープンキャプションの位置は決められており、その場合には、重ならない一定の位置とすることができる。

【0071】(2) 開始・終了タイミングのみの制御では、オープンキャプションと異なる開始・終了タイミングで字幕を表示する。これは、表示する字幕の位置を変えたくない場合に行われる。字幕の表示開始・終了タイ

ミングは、既に検出済みのオープンキャプションの開始・終了タイミングと異なるように決定する。

【0072】(3) 字幕の表示位置と開始・終了タイミングの双方の制御では、オープンキャプションと異なる位置及び開始・終了タイミングで字幕を表示する。これは、より適切に字幕を表示したい場合に行われる。字幕の表示位置及び開始・終了タイミングは、既に検出済みのオープンキャプションの表示位置及び開始・終了タイミングと異なるように決定する。

【0073】一般的には、図6に示す手順で字幕の表示位置制御を行うことができる。図6において、ステップS61では、オープンキャプション判定部39において、画面上の所定領域にオープンキャプションが存在するか否かの判定がなされる。この判定の要件は前述したが、少なくとも振幅レベルと静止時間が所定値以上あることが要件とされる。

【0074】その結果、オープンキャプションが存在すると判定された場合には(ステップS62)、表示位置決定部47及び表示タイミング決定部49では、その検出されたオープンキャプションの存在位置、大きさ、開始・終了タイミングを調査し、存在位置が表示予定字幕の表示位置と重なるか否かを判定する(ステップS63)。

【0075】次いでオープンキャプションの存在位置と表示予定字幕の表示位置とが重なる場合には、重なる時間が設定値T(秒)以下であるかどうかを判断する(ステップS64)。なお、設定値Tは、例えば2秒である。

【0076】そして、重なる時間が設定値T(秒)以上である場合には、表示位置決定部47が例えば同じフレーム内で重ならない他の位置を検索し、見つかるとその位置を表示位置に決定する(ステップS65)。これにより、オープンキャプションと異なる位置に字幕の表示位置が決定される。決定結果は表示字幕データ出力部53に渡される。

【0077】一方、重なる時間が設定値T(秒)以下である場合には、表示タイミング決定部49が例えばオープンキャプションが存在しない他の直近フレームを検索し、表示タイミングを決定する(ステップS66)。これにより、字幕の表示開始・終了タイミングが、既に検

出済みのオープンキャプションの開始・終了タイミングと異なるように決定される。決定結果は表示字幕データ出力部53に渡される。

【0078】このようにして表示位置やタイミングが決定された表示予定字幕が、表示字幕データ出力部53により、記憶装置に蓄積される(ステップS67)。なお、字幕の表示予定位置にオープンキャプションが存在しない場合(ステップS62)や存在しても重ならない場合(ステップS63)には、その旨が付された表示予定字幕が、表示字幕データ出力部53により、記憶装置に蓄積される(ステップS67)。

【0079】次に図7を用いて、図6に示す手順で実施される字幕の表示位置及び開始・終了タイミングの変更処理例を説明する。

【0080】図7において、オープンキャプションOP1、OP2と字幕J1、J2、J3が図に示すタイミングで表示されると、字幕J2、J3の表示位置では重なりが生ずる。字幕J2の重なりが例えば2秒以下であれば、変更字幕J2のように開始・終了タイミングを遅らせる。また、字幕J3の重なりが例えば2秒以上であれば、変更字幕J3のように表示位置を変更する。このようにして、オープンキャプションとの重なりが回避される。

【0081】なお、変更後の開始・終了タイミングが、カット変わりやシーンチェンジなどにより、映像・音声に対して不自然となる場合がある。また重要な映像部分との重なりなどにより映像に対して不自然となる場合がある。それらに対する処理は、実施の形態2で説明する。

【0082】(実施の形態2) 文字認識部41を用いて表示制御を行う場合は、オープンキャプション検出と字幕の表示制御に加えて、(1) オープンキャプションのテキスト化、(2) 映像のカット検出、などの結果を利用することにより、字幕表示を的確に制御できるようになる。前述したのと若干重複するが、再述する。

【0083】(1) オープンキャプションのテキスト化: オープンキャプションが検出できると、そのオープンキャプションの内容をテキスト化することが可能である。検出されたオープンキャプションは一種の図形データであるが、文字図形をテキスト化する高性能な文字図形認識ソフトが市販されており、この技術が適用可能である。

【0084】テキスト化により、オープンキャプションの内容が把握できると、オープンキャプションのタイミングに近接した近隣字幕の表示の要否、近隣字幕の一部字幕内容の削除、削減に伴うタイミングの修正など、より充実した字幕適性化が可能となる。例えば、オープンキャプションと一致した内容の字幕は表示不要であるので削減する。また、オープンキャプションで表示されている内容部分を該当字幕内容から削減するなどが行え

る。

【0085】(2)映像情報の検出による字幕表示タイミングの制御：字幕表示のタイミングに関し、考慮すべき映像の情報としては、カット、ワイプ、ディゾルブ、パン、チルト、ズームなどがある。特にカット、ワイプ、ディゾルブは、映像内容の時間・空間的な隔たりを表現する手法として多用されている。この隔たりを無視した字幕表示は不適切である場合が多い。したがって、このようなことを避ける制御が必要である。例えば、カットのタイミングを避ける字幕の開始・終了タイミングの制御などがある。

【0086】次に、図8を用いて、文字認識部41を用いた場合の表示制御動作（オープンキャプションの内容に応じた表示制御の適性化例）を説明する。

【0087】まず「(1)オープンキャプションのテキスト化」に関する処理を説明する。なお、図6と同一処理となるステップには同一符号を付してある。ここでは、異なる部分を説明する。

【0088】図8において、オープンキャプションの存在位置と表示予定字幕の表示位置とが重なる場合には（ステップS63）、文字認識部41が起動され、オープンキャプションの内容が識別される（ステップS84）。次いで、表示／非表示・修正／非修正決定部51において、オープンキャプションの内容と表示予定字幕の内容とを比較し、近似度を調査する。近似度は、2段階に渡って判断される（ステップS85、86）。

【0089】ステップS85では、近似度が閾値1よりも大きいかな否かが判断される。閾値1は、例えば近似度0.9である。近似度が高い場合には、当該字幕は表示しないと決定する（ステップS87）。また、ステップS86では、近似度が閾値2よりも大きいかな否かが判断される。閾値2は、例えば近似度0.5である。類似部分があるときは、その類似部分を削除ないしは修正する決定を行う（ステップS88）。

【0090】そして、近似度が低い場合には、図6で説明したように、重なり時間が設定値T以下であるかな否かが判断され（ステップS64）、判断結果に応じて表示位置とタイミングが制御される（ステップS65、66）。このようにして表示／非表示・修正／非修正の決定がなされた表示予定字幕が、表示字幕データ出力部53により、記憶装置に蓄積される（ステップS89）。

【0091】次に、ステップS90～S92は、「(2)映像情報の検出による字幕表示タイミングの制御」に関する処理を示している。なお、実施の形態1（図6）においても、この「映像情報の検出による字幕表示タイミングの制御」が行えることはいうまでもない。

【0092】変更後の開始・終了タイミングが、カット変わりやシーンチェンジなどにより、映像・音声に対して不自然となる場合、また重要な映像部分との重なりな

どにより映像に対して不自然となる場合があるので、映像のカット点を検出すると（ステップS90）、字幕終了タイミングと比較し（ステップS91）、不自然にならないように字幕の開始・終了タイミングを変更する（ステップS92）。このようにして開始・終了タイミングの変更決定がなされた表示予定字幕が、表示字幕データ出力部53により、記憶装置に蓄積される（ステップS89）。

【0093】次に図9を用いて、図8に示す手順で実施される字幕表示の適性化処理例を説明する。図9において、オープンキャプションOP1、OP2と字幕J1、J2、J3、J4が図に示すタイミングで表示されると、字幕J2、J3の表示位置では重なりが生ずる。また、字幕J4は、映像のカット点（図中、黒三角で示す点）に跨る。

【0094】そのため、字幕J4については、その開始・終了タイミングと映像の前記カット点とを比較し、終了タイミングを前記カット点以前となるように、開始・終了タイミングを変更する。このようにして、映像・音声との関係での不自然な表示が回避される。

【0095】また、字幕J2、J3の重なりは、オープンキャプションOP1、OP2の内容との比較で処理する。即ち、字幕J3とオープンキャプションOP2の内容が非常に近似している場合は、図に示すように、字幕J3の表示を取り止める。一方、近似していない場合は、前述したように字幕J3の表示位置を変更する。

【0096】字幕J2とオープンキャプションOP1の例では、字幕J2の内容「森首相は昨日訪米し、ブッシュ大統領と会談した。」と、オープンキャプションOP1の内容「ブッシュ大統領と会談した。」とを比較して、新しい内容「森首相は昨日訪米した。」及び開始・終了タイミングを変更した変更字幕J2'として表示する。このようにして、オープンキャプションとの重なりが適切に回避される。

【0097】なお、受信側でオープンキャプションを検出できる場合は、さらに多様な回避対策が可能である。字幕受信者が字幕を重視し、特に字幕に注視してテレビを見ているとすると、字幕は常に同じ位置に表示されているのが望ましい。字幕表示位置がその都度変わると見難いからである。また、字幕として該当内容が表示されている場合は、オープンキャプションは見えなくてもよいとも考えられる。

【0098】そのため、字幕は常に定まった位置に表示することとし、表示する字幕の内容にオープンキャプションの内容が全て、もしくは殆どが含まれる場合は、オープンキャプションを消去もしくはレベルを充分低減し、その他の場合は、オープンキャプションを所定位置の字幕表示に影響しない場所に移動して表示する処置を採ることとする。

【0099】これらの手法として例えば、消去は検出さ

れたオープンキャプション信号を再生し、映像信号から同レベルで減算する。レベル低減は、所定の比率で減算することで実現できる。また、オープンキャプションの移動は、前記の方法で消去し他の位置で映像に合成する方法による。これらの手法は、充分実現可能な技術範囲である。

【0100】

【発明の効果】以上詳細に説明したように、本発明によれば、オープンキャプションの特徴を利用してその表示位置やタイミングなどを自動的に検知し、それと重複しないように表示予定字幕の表示位置やタイミングを自動的に制御するようしたので、人手によっていた作業を自動化することができる。また、オープンキャプションの表示位置やタイミングのみならず、そのキャプションをテキストとしても自動的に検知し、表示予定の字幕内容と比較して、字幕の表示/非表示、字幕内容の修正/非修正なども自動的に判別し実行させることができるので、字幕制作の効率化が図れるようになる。したがって、今後適用分野や番組数などの拡大が見込まれる字幕放送において、このような自動化は字幕番組制作上に大きな効果が期待される。

【図面の簡単な説明】

【図1】本発明に係るオープンキャプションに対する字幕表示制御方法を実施する自動字幕番組制作システムの機能ブロック構成図である。

【図2】自動字幕番組制作システムにおける字幕制作フローを、改良された現行字幕制作フローと対比して示した説明図である。

【図3】単位字幕文を表示単位字幕文毎に分割する際に適用される分割ルールの説明に供する図である。

【図4】アナウンス音声に対する字幕送出タイミングの同期検出技術に係る説明に供する図である。

【図5】アナウンス音声に対する字幕送出タイミングの\*

\*同期検出技術に係る説明に供する図である。

【図6】本発明の実施の形態1に係るオープンキャプションに対する字幕表示制御方法を説明するフローチャートである。

【図7】実施の形態1による字幕の表示位置及び開始・終了のタイミングの変更例を示す図である。

【図8】本発明の実施の形態2に係るオープンキャプションに対する字幕表示制御方法を説明するフローチャートである。

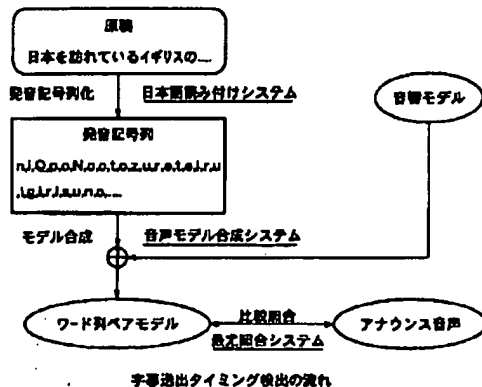
【図9】実施の形態2による字幕表示の適性化処理例を示す図である。

【図10】現行字幕制作フロー、及び改良された現行字幕制作フローを対比して示した説明図である。

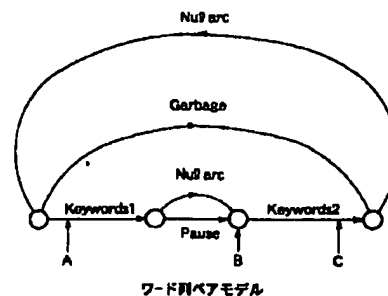
【符号の説明】

- 11 自動字幕番組制作システム
- 13 電子化原稿記録媒体
- 15 同期検出装置
- 17 統合化装置
- 19 形態素解析部
- 21 分割ルール記憶部
- 23 デジタル・ビデオ・テープ・レコーダ(D-VTR)
- 33 単位字幕文抽出部
- 35 表示単位字幕化部
- 37 タイミング情報付与部
- 39 オープンキャプション判定部
- 41 文字認識部
- 43 画面変更判定部
- 45 字幕表示位置制御部
- 47 表示位置決定部
- 49 表示タイミング決定部
- 51 表示/非表示・修正/非修正決定部
- 53 表示字幕データ出力部

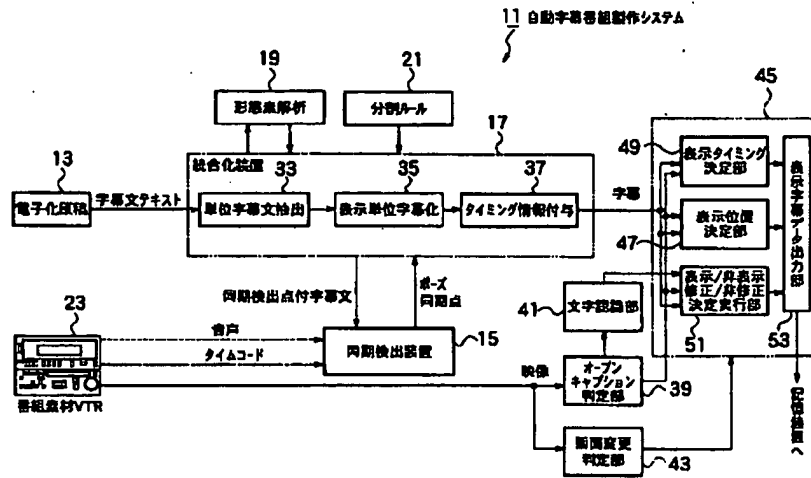
【図4】



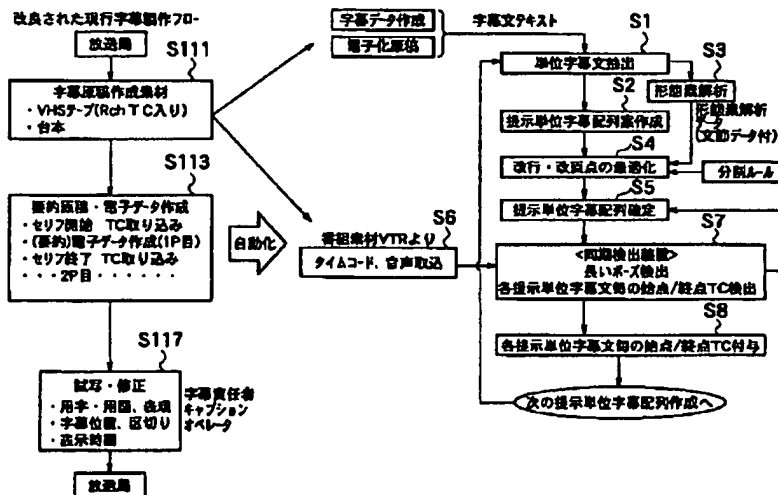
【図5】



【図1】



【図2】



【図7】

オープンキャプション	.....OP1.....	.....OP2.....
字幕	.....J1.....	.....J2.....
変更	.....J1.....	.....J2.....

【図3】

**改行・改頁ルール(分割ルール)例**

この順に逐次し、4では下記図表による。

1. 改行・改頁点は句点の後
2. 改行・改頁点は読点の後
3. 改行・改頁点は文節と文節の間
4. 改行・改頁点は形態品詞の間

下記順度の高い形態品詞の直前で改行・改頁する。

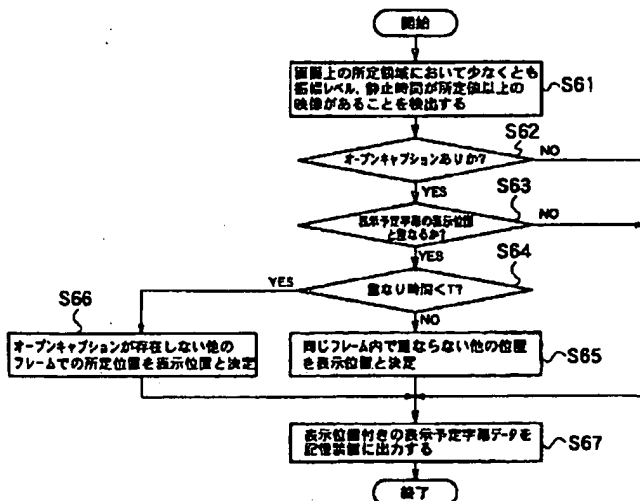
【図9】

オープンキャプション	-----OP1-----	-----OP2-----	
字幕	-----J1-----	-----J2-----	-----J3-----
映像のカット			▲
変更	-----J1-----	-----J2-----	-----J4-----

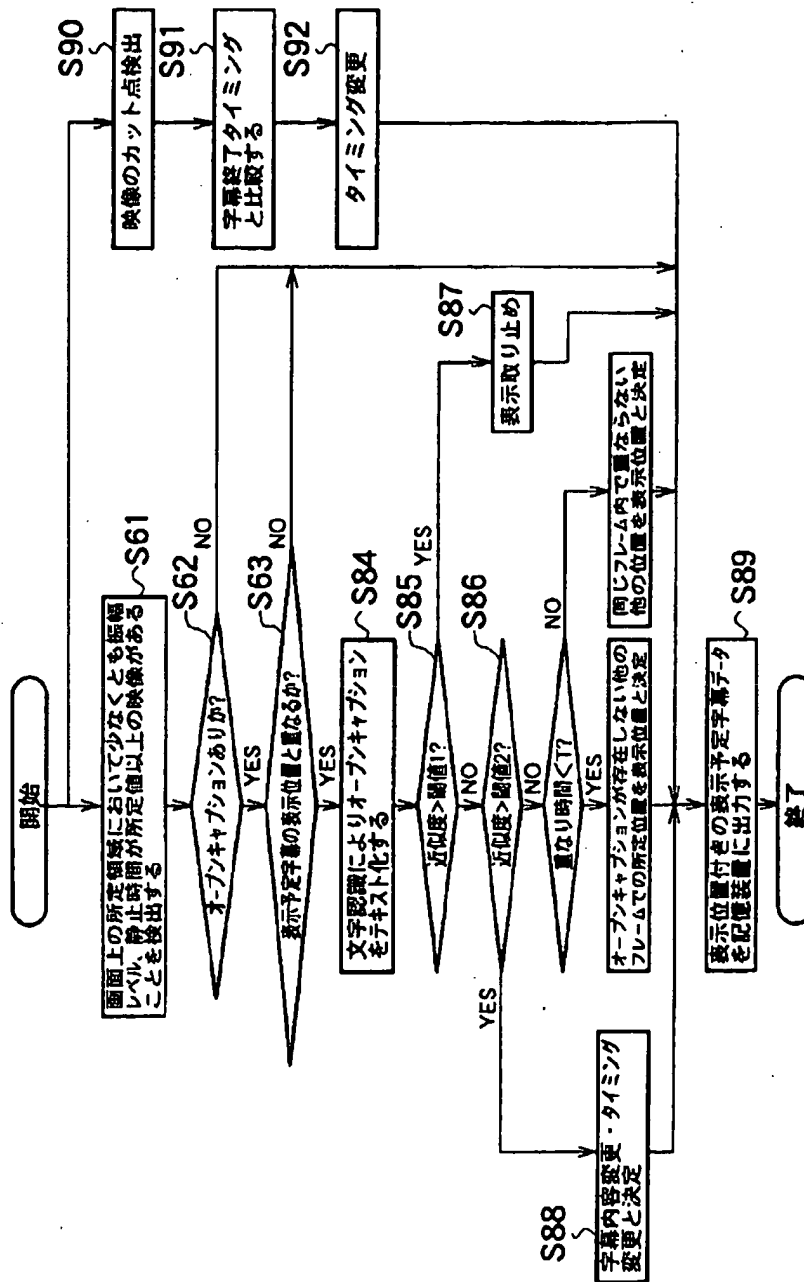
順度[形態品詞]	4[格助と]	2[助断定]	1[係助他]
12[格助に]	4[格助で]	2[格助他]	1[格助へ]
12[係助は]	4[係助]	2[助断]	1[格から]
9[格助を]	4[助]	1[助助か]	1[助名]
8[体助を]	4[助格]	1[助計名]	1[助助]
8[自属]	3[助過去]	1[助形受]	1[後助]
5[係助も]	3[助と]	1[助補助]	1[助]
5[格助が]	3[名]	1[助完了]	1[後]

自然語のある改行・改頁における、優先の形態品詞とその順度例

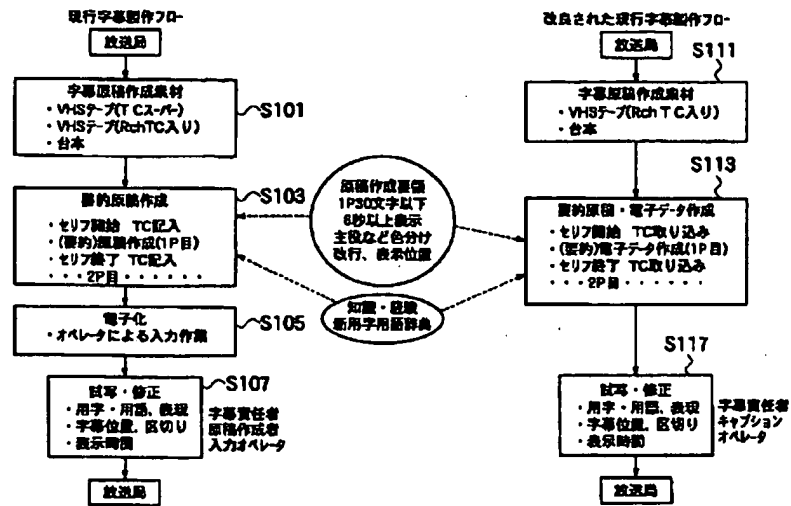
【図6】



【図8】



【図10】



フロントページの続き

(72)発明者 沢村 英治  
東京都港区芝2-31-19 通信・放送機構  
内

(72)発明者 門馬 隆雄  
東京都港区芝2-31-19 通信・放送機構  
内

(72)発明者 江原 暉将  
東京都港区芝2-31-19 通信・放送機構  
内

(72)発明者 白井 克彦  
東京都港区芝2-31-19 通信・放送機構  
内

Fターム(参考) 5C023 AA18 BA01 BA11 BA16 CA01  
CA04 CA05 EA03  
5C063 AA01 AC01 AC05 CA40 DA03  
DB09 EA03 EB01